Information Processing Letters 116 (2016) 120–124

Contents lists available at ScienceDirect

Information Processing Letters

www.elsevier.com/locate/ipl

Improving resolution width lower bounds for k-CNFs with applications to the Strong Exponential Time Hypothesis

Ilario Bonacina [a,*], Navid Talebanfard [b,1]

[a] Department of Computer Science, Sapienza University of Rome, via Salaria 113, 00198 Roma, Italy
[b] Department of Mathematical and Computing Sciences, Tokyo Institute of Technology, Meguro-ku, Ookayama 2-12-1, 152-8552, Japan

ARTICLE INFO

Article history:
Received 3 June 2014
Received in revised form 25 May 2015
Accepted 22 September 2015
Available online 13 October 2015
Communicated by A. Muscholl

Keywords:
Theory of computation
Proof complexity
Resolution
Strong Exponential Time Hypothesis

ABSTRACT

A Strong Exponential Time Hypothesis lower bound for resolution has the form $2^{(1-\epsilon_k)n}$ for some k-CNF on n variables such that $\epsilon_k \to 0$ as $k \to \infty$. For every large k we prove that there exists an unsatisfiable k-CNF formula on n variables which requires resolution width $(1 - \widetilde{O}(k^{-1/3}))n$ and hence tree-like resolution refutations of size at least $2^{(1-\widetilde{O}(k^{-1/3}))n}$. We also show that for every unsatisfiable k-CNF φ on n variables, there exists a tree-like resolution refutation of φ of size at most $2^{(1-\Omega(1/k))n}$.

1. Introduction

The *Exponential Time Hypothesis* (ETH) formulated by Impagliazzo and Paturi [6] states that the SAT problem requires exponential time. They also gave a strengthening of this, the so-called *Strong Exponential Time Hypothesis* (SETH) which states that the complexity of k-SAT grows as k increases and it approaches that of exhaustive search. More precisely let $\sigma_k = \inf\{\delta : k\text{-SAT can be solved in time } O(2^{\delta n})\}$. SETH states that $\lim_{k\to\infty} \sigma_k = 1$.

Of course these are both stronger than NP $\neq$ P and hence any proof is far beyond reach at the moment. However since the running times of the best known k-SAT algorithms have the form $2^{(1-\epsilon_k)n}$ where $\epsilon_k \to 0$ as $k \to \infty$ (see e.g. [9]), one can ask whether SETH holds for specific algorithms, that is whether there are k-CNF instances on which the algorithms run for at least $2^{(1-\epsilon_k)n}$ steps. This turns out to be the case for certain classes of algorithms. For the PPSZ algorithm such a lower bound was proved by Chen et al. [5]. For DPLL the connection with resolution complexity has been used to derive SETH lower bounds: indeed it is well known that a run of a DPLL algorithm on an unsatisfiable k-CNF gives a tree-like refutation, therefore a tree-like resolution refutation lower bound would imply a DPLL running time lower bound. Pudlák and Impagliazzo [8] constructed unsatisfiable k-CNF formulas that require tree-like resolution refutations of size $\Omega(2^{(1-\epsilon_k)n})$ where $\epsilon_k = O(1/k^{1/8})$. A recent construction by Beck and Impagliazzo [3] improves this to $\epsilon_k = \widetilde{O}(1/k^{1/4})$, where the $\widetilde{O}$ notation is hiding log factors.

In this paper we clarify and simplify the result in [3] making a further improvement to $\epsilon_k = \widetilde{O}(1/k^{1/3})$. Our main contribution is an improvement of a strong width lower bound first proved in [3], where it was shown that there are k-CNFs in n variables which require Resolution width at least $(1 - \widetilde{O}(1/k^{1/4}))n$. We improve this lower bound to $(1 - \widetilde{O}(1/k^{1/3}))n$ which combined with the

* Corresponding author.
E-mail addresses: bonacina@di.uniroma1.it (I. Bonacina), navid@is.titech.ac.jp (N. Talebanfard).

1 This work was done while N.T. was with Aarhus University. The author acknowledges support from the Danish National Research Foundation and The National Natural Science Foundation of China (under the grant 61061130540) for the Sino-Danish Center for the Theory of Interactive Computation, within which this work was performed. He also acknowledges support from the Center for Research in Foundations of Electronic Markets (CFEM), supported by the Danish Strategic Research Council.

http://dx.doi.org/10.1016/j.ipl.2015.09.013
0020-0190/

width-size relationship of Ben-Sasson and Wigderson [4] we immediately get the tree-like size lower bound. One can also ask how far this improvement can go. Using the Switching Lemma, we can show that for every unsatisfiable k-CNF on n variables, there exists a tree-like resolution of size at most $2^{(1-\Omega(1/k))n}$ (Theorem 2.2). A similar argument was used by Miltersen et al. [7] to prove upper bounds on the size of decision trees for k-CNF's. However, we are not aware of any adaptation of this in the proof complexity literature and will hence present a formal account of this observation.

Relation to [3] As already said this paper simplifies and improves the results from Beck and Impagliazzo [3]. In both papers the lower bounds are proven for a family of CNF formulas encoding unsatisfiable linear systems of equations over $\mathbb{F}_p$ with a certain kind of expansion property (Definition 3.1 and Proposition 3.2). The main technical difference is in the way such linear systems are encoded using boolean variables. Our encoding is more efficient from the point of view of the number of boolean variables used and this is essentially the reason for the improved lower bounds (Definition 3.4). The main technical improvement over [3] is Lemma 3.3. Our width lower bound (Theorem 3.5) is a modification of the analogue of Theorem 5.5 of [3]. Then, for tree-like resolution, our SETH size lower bound (Corollary 3.6) will follow immediately from the size-width relationship of [4]. In [3] is shown that SETH lower bounds hold for regular resolution, a proof system exponentially stronger than tree-like resolution. Their proof relies essentially on the properties of the linear system of equation over $\mathbb{F}_p$, which is exactly the same we use, and not on the encoding of such system as a CNF formula. Hence such proof still holds with the new encoding and our improvement from $\epsilon_k = \widetilde{O}(1/k^{1/4})$ to $\epsilon_k = \widetilde{O}(1/k^{1/3})$ still carry on for regular resolution using their proof.

Notations *Resolution* (RES) is one of the most fundamental and extensively studied proof systems. Using this proof system one can refute unsatisfiable CNF formulas using the following inference rule

$$\frac{C \vee x \quad D \vee \neg x}{C \vee D},$$

where C and D are disjunctions of literals and x is some variable. Every resolution refutation induces a DAG in the following way. There is a node for each clause appearing in the proof and every such node will be connected by an edge to the nodes corresponding to the two clauses from which this clause was derived. Given a formula φ we denote a RES refutation of φ using the notation $\varphi \vdash_\pi \bot$.

Tree-like resolution (treeRES) refers to a subclass of resolution for which the induced DAG is in fact a tree. Let C be a clause, the *width of* C, $|C|$, is the number of literals appearing in C. The resolution *width* of a formula φ denoted by $\mathrm{width}(\varphi \vdash \bot)$ is the minimum of the width of the largest clause appearing in any resolution refutation of φ, more formally

$$\mathrm{width}(\varphi \vdash \bot) := \min_\pi \max\{|C| \ : \ \varphi \vdash_\pi \bot \text{ and C appears in } \pi\}.$$

A *restriction* on a set of variables X is a mapping $\rho : X \to \{0, 1, \star\}$. We call a variable *unfixed by* ρ if it is assigned $\star$, and we call it *fixed* otherwise. The domain of ρ, denoted by $\mathrm{dom}\,\rho$, is the set of variables fixed by ρ and $|\rho| := |\mathrm{dom}(\rho)|$. For a function f, we define $f|_\rho$ to be the function after setting values to the fixed variables according to ρ. A *random restriction* leaving ℓ variables free can be obtained as follows: first pick a subset S of the variables of size $|X| - \ell$ uniformly at random, then set each $x \in S$ to either 0 or 1 with equal probability.

2. Size upper bound

A *decision tree for an unsatisfiable k-CNF* φ is a binary tree where inner nodes are labeled with variables from φ and leaves are labeled with clauses from φ. Each path in the tree corresponds to a partial assignment where a variable x gets the value 0 or 1 according whether the path branches left or right at the node labeled with x. The condition on the tree is that each clause on the leaves is falsified by the partial assignment given by the path reaching the clause. Decision trees for an unsatisfiable k-CNFs corresponds to treelike Resolution refutations (see e.g. [2]).

Following Beame [1] we define the canonical decision tree. Given a CNF $\varphi = \bigwedge_i C_i$ consider an ordering of the variables and an ordering of the clauses. The *canonical decision tree* of φ, denoted by $T(\varphi)$, is inductively defined as follows: look at the first clause C of φ according to the ordering and assume $\varphi = C \wedge \varphi'$. Then do a full decision tree on the variables of C respecting the order of the variables, i.e. along each path from the root to leaves the order on which the variables are appearing is compatible with the fixed ordering of the variables. To the leaf corresponding to the restriction which falsifies C, we assign C. For other leaves corresponding to restriction σ, we replace the leaf with $T(\varphi'|_\sigma)$. Notice that canonical decision trees in [1] are defined for general CNFs, but we can adopt them for the unsatisfiable setting. The following variant of the Switching Lemma is due to Beame [1].

Lemma 2.1 *(Switching Lemma). Let φ be a k-CNF on n variables. Let ρ be a random restriction chosen uniformly at random from the set of all restrictions that leave exactly ℓ variables unset, with $\ell \leq \frac{n}{7}$. The probability that the canonical decision tree of $\varphi|_\rho$ has depth bigger than d is at most $\left(\frac{7k\ell}{n}\right)^d$.*

Theorem 2.2. *For any unsatisfiable k-CNF φ on n variables*

$$\mathrm{size}_{\mathrm{treeRES}}(\varphi \vdash \bot) \leq 2^{(1-\Omega(\frac{1}{k}))n}.$$

Proof. We follow an argument due to Miltersen et al. [7] who showed that every k-CNF has a decision tree representation of size $2^{(1-\Omega(\frac{1}{k}))n}$ and adjust it to the unsatisfiable setting. We set $\ell = n/14k$. By the Switching Lemma, for a $1 - 2^{-d}$ fraction of restrictions σ with $|\sigma| = n - \ell$, we know that the depth of $T(\varphi|_\sigma)$ is at most d.

Then, by an averaging argument, there exists a subset S of the variables of φ with $|S| = n - \ell$ such that the same statement holds for all restrictions fixing variables only in S, that is for at least $1-2^{-d}$ of the assignments σ setting exactly the variables in S, $T(\varphi|_\sigma)$ is at most d.

We can construct a decision tree for φ as follows: first we do a full decision tree on variables in S; then for each leaf with the corresponding restriction σ, we append $T(\varphi|_\sigma)$ to that leaf. The number of leaves of this tree is upper bounded by

$$2^d 2^{n-\ell} + 2^{-d} 2^{n-\ell} 2^{\ell},$$

since at most a 2^{-d} fraction of the leaves of the full decision tree on S can have maximal depth 2^ℓ. Setting $d := \ell/2$ the number of leaves of the tree is upper bounded by $2^{n-\frac{\ell}{2}+1} = 2^{(1-\Omega(\frac{1}{k}))n}$. Hence the required upper bound on the size of the decision tree constructed. Since decision trees correspond to treelike resolution refutations we have the desired upper bound on the size of treelike refutations of φ. □

3. Size lower bound

Let $\mathbf{v} = (v_1, v_2, \ldots)$ be a vector over $\mathbb{F}_p$, then by $\mathrm{supp}(\mathbf{v})$ we denote the indices of $\mathbf{v}$ with non-zero entries, that is $\mathrm{supp}(\mathbf{v}) := \{i \ : \ v_i \neq 0 \mod p\}$.

In what follows we construct a system of linear equations over $\mathbb{F}_p$. Let m be the total number of equations, n the total number of variables. We use the letter E, with subscripts, to denote linear equations mod p, that is expressions of the form

$$\sum_j a_j z_j = b \mod p,$$

with $a_j, b \in \mathbb{F}_p$. We denote with $\mathrm{supp}(E)$ the set of indices j having non-zero a_js.

Definition 3.1 *((α,β,γ)-expander).* Let $\alpha,\beta,\gamma \in \mathbb{R}_{\geq 0}$, $m, n \in \mathbb{N}$ and $\mathscr{E} := \{E_1, \ldots, E_m\}$, a set of m linear equations over $\mathbb{F}_p$. We say that $\mathscr{E}$ is an *(α,β,γ)-expander* if and only if

$$\forall \mathbf{v} \in \mathbb{F}_p^m,\ \alpha \leq |\mathrm{supp}(\mathbf{v})| \leq \beta \Rightarrow \left|\mathrm{supp}\left(\sum_{i=1}^m v_i E_i\right)\right| \geq \gamma.$$

The previous definition is essentially the same from [3] and the next proposition is a particular case of Lemma 4.2 from [3].

Proposition 3.2. *Let p a sufficiently large prime. There exists a set $\mathscr{E} := \{E_1, \ldots, E_{n+1}\}$ of linear equations in n variables over $\mathbb{F}_p$ such that:*

1. *$\mathscr{E}$ is unsatisfiable,*
2. *for each $E_i \in \mathscr{E}$ $|\mathrm{supp}(E_i)| \leq p^2$,*
3. *$\mathscr{E}$ is $(\delta n, 3\delta n, (1-c\theta)n)$-expander, where $\delta = O(1/p)$, $\theta = \widetilde{O}(1/p)$ and c is a constant,*
4. *no subset of at most $3\delta n$ equations from $\mathscr{E}$ is unsatisfiable.*

In [3] they encode each variable of the set of linear equations from Proposition 3.2 using a sum of roughly p^2 boolean variables and show that with this encoding the linear system requires very large resolution width.

The key property of this representation used in the proof is the following: let $z = \sum_{i=0}^{p^2} x_i$, where x_i are boolean variables. Even setting a lot of variables (i.e. $p^2 - p$) we still can obtain all possible $\mathbb{F}_p$ values for z setting the remaining variables.

In other words what Beck and Impagliazzo really require is a function that can extract $\log p$ bits even after many bits in the input are fixed. Our contribution is thus to show that a random function satisfies this property (Lemma 3.3), and we use this function instead of the sum of p^2 boolean variables used by Beck and Impagliazzo. The arguments of [3] still goes through (Theorem 3.5). Beck and Impagliazzo use roughly p^2 bits for each $\mathbb{F}_p$ variables, whereas with our construction we only require around p bits and hence we get the improvement.

The following lemma is then the main technical improvement over the construction in [3].

Lemma 3.3. *Let θ and p be the parameters coming from Proposition 3.2. Then there exists a function $g : \{0,1\}^{\theta^{-1}\log^2 p} \to \{0,1\}^{\log p}$ such that for any restriction σ with $|\sigma| \leq \theta^{-1}\log^2 p - \log^2 p$ we have $\mathrm{Img}(g|_\sigma) = \{0,1\}^{\log p}$.*

Proof. Let $u := \theta^{-1}\log^2 p$ and g be random function that assigns to every $x \in \{0,1\}^u$ a value in $\{0,1\}^{\log p}$ independently and uniformly at random. We bound the probability that there exist a $y \in \{0,1\}^{\log p}$ and a restriction σ with $|\sigma| = u - \log^2 p$ such that $y \notin \mathrm{Img}(g|_\sigma)$. This is easily given as follows

$$\begin{aligned}&\Pr[\exists y, \sigma \ : \ y \in \{0,1\}^{\log p}, |\sigma| = u - \log^2 p, y \notin \mathrm{Img}(g|_\sigma)] \\ &\quad \leq p \binom{u}{\log^2 p} 2^{u-\log^2 p} (1-1/p)^{2^{\log^2 p}} = o(1),\end{aligned}$$

since $u = \widetilde{O}(p)$. Then clearly we have that there must exist at least one function g realizing the complementary event that we bounded. Such function works also for each σ such that $|\sigma| \leq u - \log^2 p$. □

Let $\{z_1, \ldots, z_n\}$ be a set of variables taking values over $\mathbb{F}_p$. The function $g : \{0,1\}^{\theta^{-1}\log^2 p} \to \{0,1\}^{\log p}$ obtained from Lemma 3.3 can be used to define each variable z_i over $\mathbb{F}_p$ using $u = \theta^{-1}\log p$ new boolean variables $x_{i1}, \ldots, x_{iu}$:

$$z_i = \sum_{j=1}^{\log p} 2^{j-1} g_j(x_{i1}, \ldots, x_{iu}), \tag{1}$$

where g_j represents the j-th coordinate of g. Hence a linear equation mod p in n variables, say

$$\sum_i a_i z_i = b \mod p,$$

can be transformed into a boolean function using equation (1) using $N := nu = n\theta^{-1}\log^2 p$ boolean variables x_{ij}:

$$\sum_{j=1}^{n} a_{ij} \sum_{k=1}^{\log p} 2^{k-1} g_k(x_{i1}, \ldots, x_{iu}) = b_i \mod p$$

Moreover if $|\mathrm{supp}(a_1,\dots,a_n)| \le d$ then the boolean encoding of this function as a CNF turns out to be a (du)-CNF. More precisely we have the following definition.

Definition 3.4. Take the set $\mathscr{E} := \{E_1,\dots,E_{n+1}\}$ of linear equations in n variables over $\mathbb{F}_p$ from Proposition 3.2. Let E_i be the linear equation $\sum_{j=1}^{n} a_{ij} z_j = b_i \mod p$ with $a_{ij}, b_i \in \mathbb{F}_p$. Replacing each z_j with the expression given in (1) we obtain a boolean function

$$E_i^b := \sum_{j=1}^{n} a_{ij} \sum_{k=1}^{\log p} 2^{k-1} g_k(x_{i1},\dots,x_{iu}) = b_i \mod p.$$

The CNF formula φ we will consider is the encoding of the following boolean function:

$$\varphi := \bigwedge_{i=1}^{m} E_i^b.$$

Note that since for each i we have $|\mathrm{supp}(E_i)| \le p^2$, φ is a $(p^2\theta^{-1}\log^2 p)$-CNF in $N := n\theta^{-1}\log^2 p$ variables.

Let $\mathscr{E}^b := \{E^b : E \in \mathscr{E}\}$ and $\mu(C) := \min\{|S| : S \subseteq \mathscr{E}^b \wedge S \models C\}$. We say that a clause C has *medium complexity w.r.t.* μ iff $\mu(C) \in (\frac{3}{2}\delta n, 3\delta n]$, with δ the parameter coming from Proposition 3.2.

The proof of the following theorem is similar to the analogous result in [3].

Theorem 3.5. *Let φ be the unsatisfiable CNF coming from Definition 3.4, μ as above and let C be a clause over the x_{ij} variables of medium complexity w.r.t. μ then*

$$\mathrm{width}(C) \ge (1-(c+1)\theta)N,$$

where c and θ are as in Proposition 3.2.

Proof. Let C be a clause of medium complexity, that is $\mu(C) \in (\frac{3}{2}\delta n, 3\delta n]$ and by contradiction $\mathrm{width}(C) < (1-(c+1)\theta)N$. Take the minimal restriction ρ setting C to $\bot$, then $|\rho| = \mathrm{width}(C)$.

We say that a variable z_i is *free* if and only if

$$|\mathrm{dom}\,\rho \cap \{x_{i1},\dots,x_{iu}\}| \le u - \log^2 p.$$

First we prove that there are at least $c\theta n$ free variables.

Let Z be the number of z_i variables that are free. We have an upper bound for the number of x_{ij} variables non-assigned by ρ:

$$(c+1)\theta N < N - \mathrm{width}(C) \le (n-Z)(\log^2 p) + uZ.$$

Hence

$$c\theta N + \theta N < n\log^2 p - Z\log^2 p + uZ.$$

Now if $Z \le c\theta n$ a contradiction follows immediately recalling that $N = un$ and $\theta N = n\log^2 p$.

An extension of ρ to all the x_{ij} variables for i such that z_i is not free induces a restriction over the z_i variables mapping them in $\mathbb{F}_p$: let ρ^* denote such an extension. We look at it both as a restriction over the x_{ij} variables or a restriction (taking values in $\mathbb{F}_p$) over the z_i variables.

So the z_i variables that are free are exactly, by construction, the ones unfixed by ρ^*. As observed we have that the number of free variables is at least $c\theta n$ and hence $|\rho^*| \le n - c\theta n = (1-c\theta)n$.

As C is of medium complexity, there exists some set of equations $S \subseteq \mathscr{E}^b$ such that $S \models C$, $|S| \in (\frac{3}{2}\delta n, 3\delta n]$ and S is minimal w.r.t. inclusion.

This implies that for each possible ρ^* of the form described above, $\{S|_{\rho^*}\}$ is unsatisfiable. Moreover, by minimality of S, for each equation $E \in S$ there exists some ρ^* such that $E|_{\rho^*}$ is not an empty constraint.

The fact that, for each ρ^* we have that $S|_{\rho^*}$ is unsatisfiable means exactly that for all ρ^* there exists some $v \in \mathbb{F}_p^m$ (dependent on ρ^*) with $|\mathrm{supp}(v)| \le |S|$ and $\sum_i v_i E_i|_{\rho^*}$ is unsatisfiable. Hence for each such ρ^* $\mathrm{supp}(\sum_i v_i E_i) \subseteq \mathrm{dom}(\rho^*)$, otherwise we could use the variables unfixed by ρ^* to satisfy $\sum_i v_i E_i$. Let $E^{\rho^*} := \sum_i v_i E_i|_{\rho^*}$ (where v depends on ρ^*).

Take a random linear combination of all the E^{ρ^*} for all the possible ρ^*: $\sum_{\rho^*} \alpha_{\rho^*} E^{\rho^*}$. Again we have that $\mathrm{supp}(\sum \alpha_{\rho^*} E^{\rho^*}) \subseteq \bigcup_{\rho^*} \mathrm{dom}(\rho^*)$.

Each E_i from S appears in this sum and its coefficient is uniformly random, and hence by averaging, there exists a linear combination such that at least $(1-1/p)\frac{3}{2}\delta n \ge \delta n$ of the E_i have non-zero coefficient. But this contradicts the expansion property as we have that $|\mathrm{supp}(\sum_{\rho^*} \alpha_{\rho^*} E^{\rho^*})| \le |\bigcup_{\rho^*} \mathrm{dom}(\rho^*)| = (1-c\theta)n$. $\square$

Corollary 3.6. *For any large enough $k \in \mathbb{N}$ there exists an unsatisfiable CNF φ in N variables such that*

- *φ is a k-CNF;*
- $\mathrm{size}_{\mathrm{treeRES}}(\varphi \vdash \bot) \ge 2^{(1-\widetilde{O}(k^{-1/3}))N}$.

Proof. Let φ be the unsatisfiable CNF formula coming from Definition 3.4. Recall that φ is a $(p^2\theta^{-1}\log^2 p)$-CNF in $N := n\theta^{-1}\log^2 p$ variables, where $\theta = \widetilde{O}(1/p)$. We have that if $C, D \models E$ then $\mu(E) \le \mu(C) + \mu(D)$ and hence in each possible refutation of φ there will be a clause of medium complexity. Hence from the previous Theorem we have that $\mathrm{width}(\varphi \vdash \bot) \ge (1-\widetilde{O}(1/p))N$. Hence $\mathrm{width}(\varphi \vdash \bot) = (1-\widetilde{O}(k^{-1/3}))N$ where $k = p^2\theta^{-1}\log^2 p$ is the width of φ. Then by the width-size relationship from [4] we have the size lower bound for treelike Resolution. $\square$

Acknowledgements

We are grateful to Albert Atserias and Nicola Galesi for discussions and suggestions. Part of this work was done while the authors were visiting KTH Royal Institute of Technology. We thank Jakob Nordström for his hospitality and for discussions and for providing an inspiring environment.

References

[1] Paul Beame, A switching lemma primer, Technical report, Department of Computer Science and Engineering, University of Washington, 1994, UW-CSE-95-07-01.

[2] Olaf Beyersdorff, Nicola Galesi, Massimo Lauria, A characterization of tree-like Resolution size, Inf. Process. Lett. 113 (18) (2013) 666–671.

[3] Christopher Beck, Russell Impagliazzo, Strong ETH holds for regular resolution, in: Proceedings of the Forty-Fifth Annual ACM Symposium on Theory of Computing, STOC '13, ACM, 2013, pp. 487–494.

[4] Eli Ben-Sasson, Avi Wigderson, Short proofs are narrow – resolution made simple, J. ACM 48 (2) (2001) 149–169.

[5] Shiteng Chen, Dominik Scheder, Navid Talebanfard, Bangsheng Tang, Exponential lower bounds for the PPSZ k-SAT algorithm, in: SODA, 2013, pp. 1253–1263.

[6] Russell Impagliazzo, Ramamohan Paturi, On the complexity of k-SAT, J. Comput. Syst. Sci. 62 (2) (2001) 367–375.

[7] Peter Bro Miltersen, Jaikumar Radhakrishnan, Ingo Wegener, On converting CNF to DNF, Theor. Comput. Sci. 347 (1–2) (2005) 325–335.

[8] Pavel Pudlák, Russell Impagliazzo, A lower bound for DLL algorithms for k-SAT (preliminary version), in: SODA, 2000, pp. 128–136.

[9] Ramamohan Paturi, Pavel Pudlák, Michael E. Saks, Francis Zane, An improved exponential-time algorithm for k-SAT, J. ACM 52 (3) (2005) 337–364.